

Introduction

An electrocardiogram (ECG) is a recording of the electricity of the heart. Analysis of ECG data can give important information about the health of the heart and can help physicians to diagnose cardiac arrhythmias, acute myocardial infarctions, conduction abnormalities, and many other heart diseases. ECGs can also be used to determine heart rate by calculating the time between successive QRS complexes. It is important to be able to calculate the heart rate between every beat since this makes it possible to look at the beat-to-beat variability in heart rate. The purpose of this laboratory was to find a reliable way to use an ECG signal to calculate the heart rate and look at heart rate variability.

This variability is of great clinical importance. In healthy individuals, there is period variation in heart rate that correlates with breathing. This variability can be explained by increased vagal stimulation to the sinoatrial node during inhalation and decreased stimulation during exhalation. Decreased heart rate variability is therefore used as a clinical marker of reduced vagal activity. Further, reduced heart rate variability has been shown to predict sudden death in patients with myocardial infarction.

In this laboratory, we varied sampling rates, analog filters, and digital filters and looked at how these variables affected the QRS peaks. Finally, we wrote four different algorithms in MATLAB for detecting the QRS peaks, and we compared their resilience to different types of noise.

Materials and Methods

We collected ECG data from multiple subjects in a variety of conditions so that we would be sure that our method of calculating heart rate variability was reliable and robust. We collected ECG data using the setup in Figure 1.

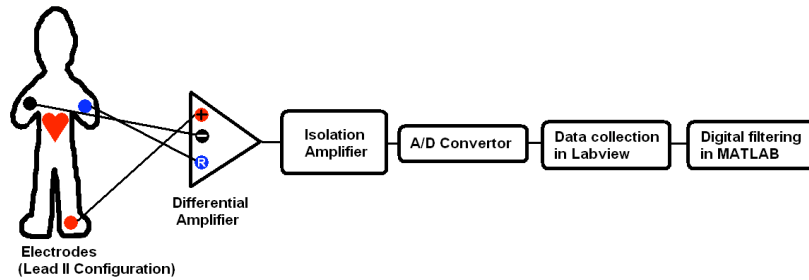


Figure 1: Our laboratory setup for collecting and analyzing ECG data

We put silver/silver chloride electrodes on skin that was cleaned and slightly abraded to reduce skin impedance. We arranged our electrodes in Einthoven's triangle to read the Lead II ECG signal because QRS complex has the largest projection on the axis of this lead. We wrapped the negative and positive leads together to shield the signal from 60 Hz noise.

Next, we passed our signal through a differential amplifier (Iso-Dam-C). The differential amplifier reduces noise common to both leads. The amplifier also has a very high input impedance (calculated to be $1.5 \text{ M}\Omega$ with a voltage divider circuit), which makes the skin impedance negligible. We set the amplifier gain to 1,000 because the amplitude of an ECG signal is generally less than 1 mV. This allowed us to pass the signal with maximum gain without worrying about saturation. We used analog filters in the differential amplifier to do a preliminary filter of the signal. Although we experimented with analog filtering, the majority of our data was collected with a low pass filter set at 200 Hz and a high pass filter set at 0.1 Hz. These filter settings were chosen to filter out high frequency noise and baseline drift respectively. After the differential amplifier, we passed the signal through an isolation amplifier (Iso-122) so that there was no direct contact between our subject and the wall socket.

We sampled the data with an A/D converter (Data Translations DT9804), and collected the signal in Labview. We experimented with different sampling rates, but we sampled all of our data at least 300 Hz in order to prevent aliasing.

Finally, we analyzed our data in MATLAB. We experimented with different types of digital filtering, different ways to calculate power spectra, and different algorithms for detecting the QRS complexes and calculating heart rate.

Results

We first investigated aliasing, which is the distortion of frequency introduced by an inadequately sampled signal. We varied the frequency of a sinusoidal analog signal from 0 Hz to 200 Hz, sampled the signal at 100 Hz, and measured the apparent frequency of the signal. This data is plotted in Figure 2.

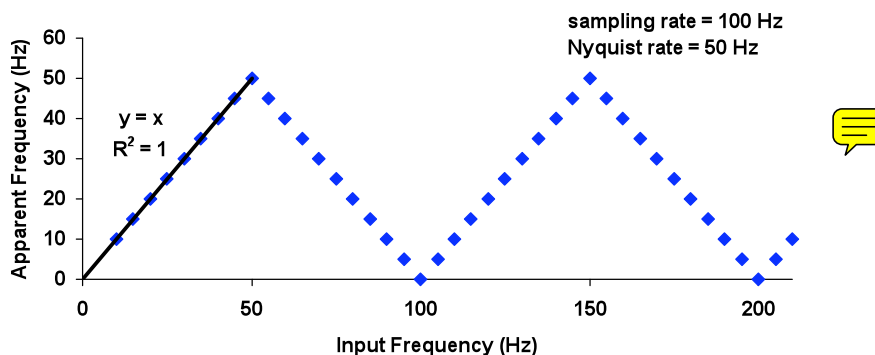


Figure 2: Plot of apparent frequency sampled at 100 Hz versus input frequency. Note that the apparent frequency is equal to the input frequency for all values below the Nyquist rate. The R^2 value for this model is 1.

From this plot, we see that the apparent frequency is equal to the input frequency for all values of frequency less than the Nyquist rate. Frequency values above the Nyquist rate are aliased into the 0 to 50 Hz region.

This means that in order to collect a good ECG signal, the sampling rate must be more than double the highest non-noise frequency component of the ECG signal. To find this highest frequency component, we looked at the power spectrum of an ECG signal. From looking at the ECG signal on an oscilloscope, we saw that the frequency of the highest frequency component of an ECG wave (the QRS complex) is approximately 20 Hz. We therefore sampled the ECG at 1 kHz because we knew that would be way more than double the highest useful frequency component.

We calculated the power spectrum using the Welch method. We wanted a frequency resolution of 1 Hz, and we wanted to calculate the power spectrum over 60 overlapping window to reduce noise, so we collected one minute of data. The power spectrum is shown in Figure 3:

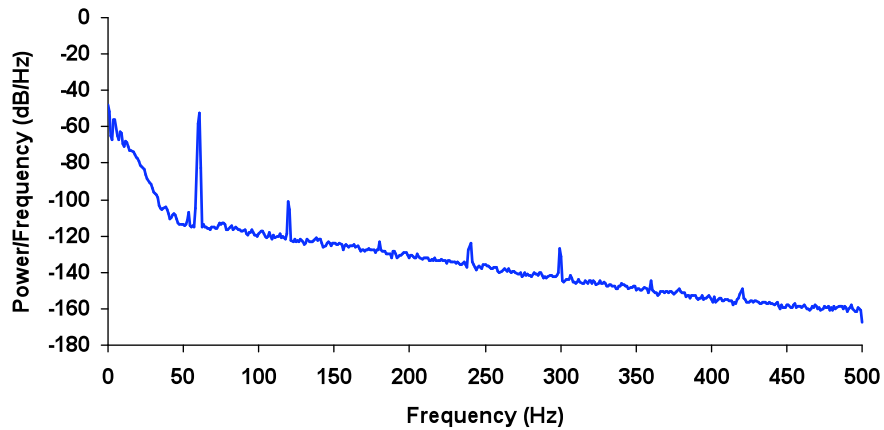


Figure 3: Plot of the power spectrum of an ECG. The sampling rate was 1000 Hz and the Nfft was 1028, which gives a frequency resolution of approximately 1 Hz. Note that excluding the 60 Hz noise and its harmonics, the high power frequencies in the ECG are all above 50 Hz.

From this power spectrum, we see that except for the 60 Hz noise and its harmonics, the high power frequencies of the ECG signal are all above 50 Hz. This means that at an absolute minimum, we must use a sampling rate of 100 Hz. However, in order to get a cleaner, better constructed signal, we sampled all of our data at 1 kHz.

Next, we examined different digital filters and their effects on the ECG signal. Our goal in digital filtering was to attenuate everything except for the QRS peaks. Specifically, we wanted to attenuate the 60 Hz noise, other high frequency noise, and motion artifacts from breathing and other movement. To attenuate the 60 Hz noise and other high frequency noise, we passed the signal through a 20 Hz, 20th order, low pass Butterworth filter. To attenuate the motion artifacts and baseline drift, we passed the signal through a 2 Hz, 20th order, high pass Butterworth filter. We chose to use 20th order filters because they have a relatively fast drop off without the risk of being unstable. We chose to use a Butterworth filters because there is no ripple in the pass band so there

is not a lot of distortion. Although these filters changed the shape of ECG signal somewhat, these filters are useful for our application because they accentuate the QRS complexes in comparison to the rest of the data. Figure 4 below shows a 6 second piece of raw ECG data, and Figure 5 below shows this same piece of ECG data that has been passed through the above filters.

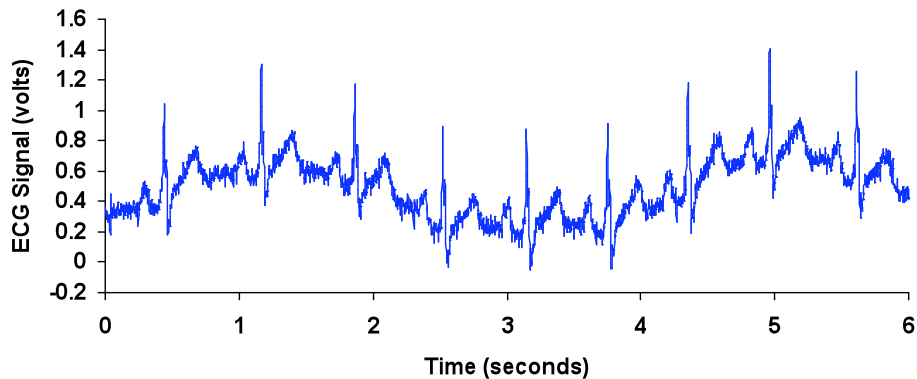


Figure 4: Plot of raw ECG data versus time. Note the presence of motion artifacts due to respiration, 60 Hz noise, and other high frequency noise.

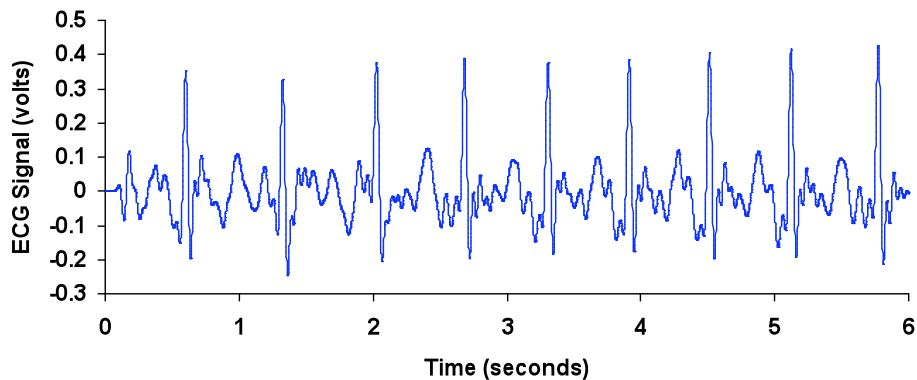


Figure 5: Plot of filtered ECG data versus time. This data was filtered with a 20 Hz, 20th order, low pass Butterworth filter to attenuate 60 Hz and other high frequency noise. It was then filtered with a 2 Hz, 20th order, high pass Butterworth filter to attenuate the motion artifacts and baseline drift.

Next, we examined various algorithms for detecting heart rate. We wrote three different algorithms in MATLAB. Algorithm 1 detects QRS peaks by setting a condition that both the amplitude and the slope of a QRS peak candidate must exceed thresholds. Algorithm 2 detects QRS peaks by setting a condition that both the slope and the inflection of a QRS peak candidate must

exceed thresholds. Algorithm 3 detects QRS peaks by setting a condition that the square of the slope of a QRS peak candidate must exceed a threshold.

When we tested these algorithms with data that was digitally filtered to attenuate everything except the QRS peaks as described above, they all worked flawlessly. Therefore, we tested each algorithm with different types of unfiltered data. The first type is “noisy” data, where we did not filter out any noise, but we asked our subject to remain as still as possible to reduce motion artifacts. The second type is “motion” data, where we asked our subject to move around slowly and we filtered out all high frequency noise with a 20th order, 30 Hz low pass Butterworth filter. Figure 6 shows how well each of our algorithms worked on “noisy” data, and Figure 7 shows how well each of our algorithms worked on “movement” data. The green diamonds are peaks that Algorithm 1 detected, the orange diamonds are peaks that Algorithm 2 detected, and the red diamonds are peaks that Algorithm 3 detected. The percent of QRS peaks detected and the percent of false positives for each algorithm for each type of data is presented in Table 1.

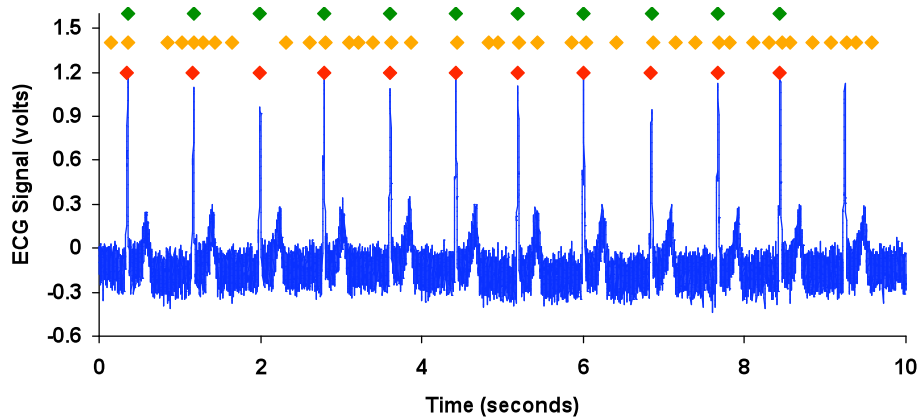


Figure 7: Plot of “noisy” ECG data versus time. The green dots are peaks that Algorithm 1 detected, orange dots are peaks that Algorithm 2 detected, and red dots are peaks that Algorithm 3 detected. Note that Algorithms 1 and 3 worked perfectly, but Algorithm 2 did not.

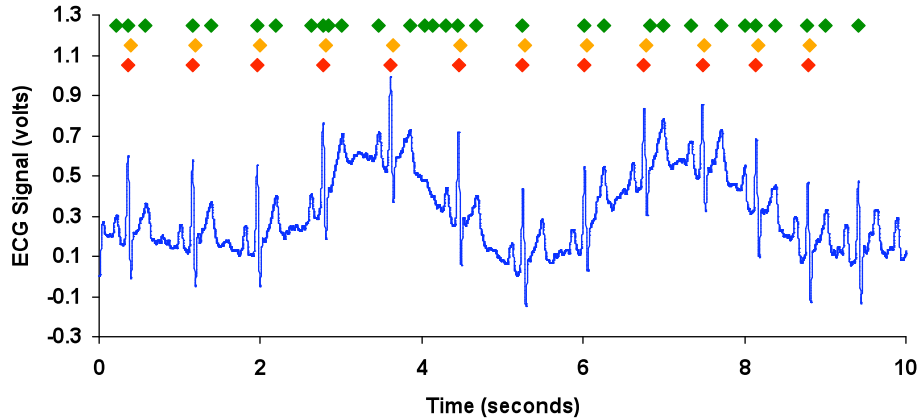


Figure 8: Plot of “movement” ECG data versus time. The green diamonds are peaks that Algorithm 1 detected, orange diamonds are peaks that Algorithm 2 detected, and red diamonds are peaks that Algorithm 3 detected. Note that Algorithms 2 and 3 worked perfectly, but Algorithm 1 did not.

	Noisy Data		Movement Data	
	% QRS Complexes Detected	% False Positives	% QRS Complexes Detected	% False Positives
Algorithm 1	100	0	83	68
Algorithm 2	91	74	100	0
Algorithm 3	100	0	100	0

Table 1: This table shows the success of each of the algorithms on “noisy” data and “movement” data. Note that Algorithm 3 works well on both types of data, Algorithm 2 works well only on “movement” data, and Algorithm 3 works well only on “noisy” data.

Discussion

In this laboratory, we found that the collection and analysis of ECG data must be catered to a specific application. For example, if we were trying to collect ECG data to look at the shape of the signal to diagnose conduction abnormalities, we would have sampled at a very high rate so that the shape would not be distorted at all. We would have then used spike triggered averaging to reduce noise across the whole spectrum. Alternatively, we could have used an FIR notch filter to pass the frequencies of interest without distorting the shape of the wave.

However, in this laboratory, we were not interested in the shape of the ECG. We were only interested in the heart rate information that can be extracted from the ECG. There are two ways to

get information about the heart rate from the ECG signal: looking at the power spectrum and detecting the QRS complexes.

It is impractical to get information about the heart rate from the ECG power spectrum. In a healthy subject, there is a characteristic high frequency heart rate component between 0.18Hz and 0.4 Hz that is synchronous with breathing. There is also a characteristic low frequency component between 0.04 Hz and 0.15 Hz that appears to be mediated by the vagus nerve¹. In order to distinguish these frequency components, we would need a frequency resolution of at least 0.01 Hz. If we sampled our ECG at 1 kHz to get a clean signal, and we wanted to average the power spectrum over 20 windows to reduce the noise, we would need to collect 33 minutes of data. We would then have 2×10^6 data points, so we would need a considerable amount of computing power to calculate the power spectrum. This method is not practical in a clinical situation.

Therefore, in this laboratory, we focused on detecting the time between QRS peaks as a method for looking at heart rate. In order to more easily see the QRS peaks in our data, we used a 20th order, notch Butterworth filter with the low frequency cutoff at 2 Hz to get rid of motion artifacts and the high frequency cutoff at 20 Hz to get rid of 60 Hz noise and other high frequency noise. With this method of filtering, the ECG wave was very distorted and of almost no clinical value for diagnosing conduction abnormalities. However, the QRS peaks are very clear in data filtered this way. All of our QRS peak detection algorithms worked flawlessly with this kind of filtered data.

We tested our algorithms on different types of unfiltered data because they all worked the same on filtered data. We found that all of the algorithms were still very good at detecting QRS peaks—at least 90% of the QRS peaks were detected by all three algorithms for both types of noise. Therefore, we were most concerned by the number of false positives that each algorithm detected.

¹ Stewart, Judith. "Heart Rate Variability." Research Network on Socioeconomic Status and Health. 1999. University of California at San Francisco. 28 Oct. 2007 <<http://www.macses.ucsf.edu/Research/Allostatic/notebook/heart.rate.html>>.

Algorithm 1 detects QRS peaks by setting a condition that both the amplitude and the slope of a QRS peak candidate must exceed thresholds. This algorithm works very well on data with 60 Hz and other high frequency noise. This is because the amplitude of the 60 Hz and other high frequency noise on the baseline or on the P or T waves is not high enough to meet the amplitude threshold, so there are no false positives from this type of noise. However, this algorithm does not work as well when there are motion artifacts. The presence of motion artifacts allows the peaks of some P and T waves to exceed the amplitude threshold, so there are a lot of false positives.

Algorithm 2 detects QRS peaks by setting a condition that both the slope and the inflection of a QRS peak candidate must exceed thresholds. This algorithm does not work as well on data with 60 Hz and other high frequency noise. This is because the slope and inflection of the noise sometimes exceeds the thresholds. When the thresholds are increased to accommodate for this, the algorithm fails to detect all of the QRS peaks. However, this algorithm works very well when there are motion artifacts. Once the 60 Hz and other high frequency noise are removed, the only peaks that are detected by this algorithm are the true QRS peaks.

Algorithm 3 detects QRS peaks by setting a condition that the square of the slope of a QRS peak candidate must exceed a threshold. This algorithm works very well for both noisy data and data with a motion artifact. This is because when slope of the ECG is squared, this accentuates the QRS peaks and attenuates everything else, which is similar to passing the data through the notch filter described above.

We determined that the best way to find the heart rate from an ECG signal is to first collect data with analog filters notch filters with a low frequency cutoff of 0.1 Hz and a high frequency cutoff of 200 Hz. The low frequency filter helps to remove baseline drift and any DC offset. The 200 Hz filter helps to eliminate aliasing of high frequency noise into the region of interest when the data is sampled. We sampled the data at 1 kHz in order to get a clean signal. Next, we digitally

filtered the data with a 20th order, notch Butterworth filter with a low frequency cutoff of 2 Hz and a high frequency cutoff of 20 Hz. This filter removed motion artifacts and 60 Hz noise, which in affect accentuated the QRS complexes and attenuated everything else. In order to detect the QRS peak, we used an algorithm that sets a condition that the square of the slope of a QRS peak candidate must exceed a threshold.

Once a reliable way to find heart rate has been determined, there are many applications. Perhaps most importantly, it is possible to look at the heart rate variability of a resting heart. To look at heart rate variability, we collected one minute of data from a resting subject, digitally filtered the data, and detected the QRS complexes with our algorithm. We then plotted heart rate versus time. Figure 9 shows the heart rate versus time for three different subjects.

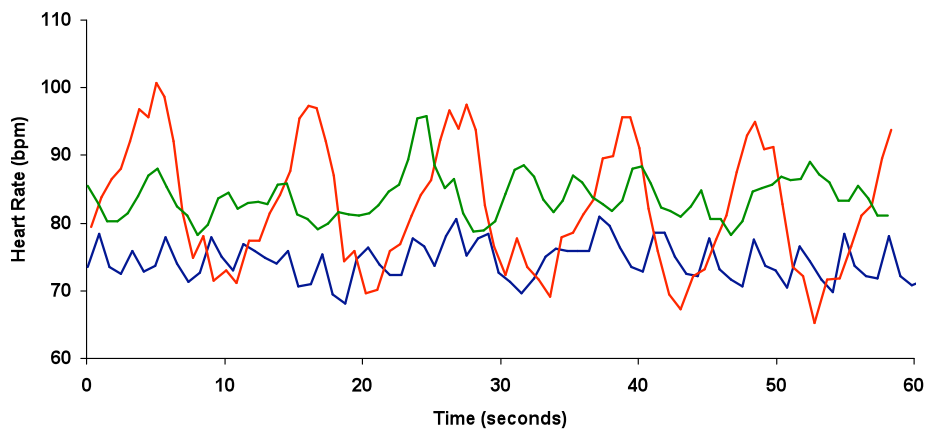


Figure 9: Plot of resting heart rate versus time for three different subjects. Note that there is a periodic variability of heart rate in time, and this variability is different for each subject. Subject 1’s trace is in red, Subject 2’s trace is in green, and Subject 3’s trace is in blue.

From this data, we found that the heart rate variability (the standard deviation of the heart rate over time) for each subject was different. Subject 1’s resting heart rate variability was 9.3 beats per minute, Subject 2’s resting heart rate variability was 3.4 beats per minute, and Subject 3’s heart resting heart rate variability was 2.9.

This heart rate variability has been shown to be correlated with breathing. It would be an interesting further experiment to detect the peaks in the heart rate versus time plot to calculate the respiration rate and compare this value with the actual respiration rate. From the plot, it looks as though Subject 1 is taking 6 breaths per minute, Subject 2 is taking 11 breaths per minute, and Subject 3 is taking 18 breaths per minute.

Another interesting application of extracting the heart rate from the ECG is looking at how heart rate recovers from exercise. We asked our three subjects to run up and down six flights of stairs, and then we recorded their ECG for the next 70 seconds. We digitally filtered the data, detected the QRS complexes with our algorithm, and plotted heart rate versus time. It has been shown that heart rate recovers parabolically with time², so we fit our data to parabolic curves. This data and the models are shown in Figure 10.

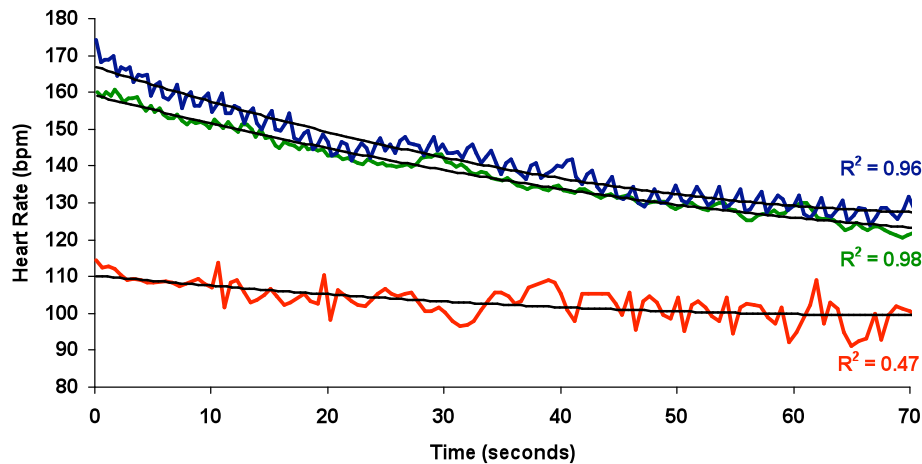


Figure 10: Plot of recovery heart rate versus time for three different subjects. The subjects were each asked to go up and down six flights of stairs. The data has been fitted to parabolic curves, and the R^2 values are shown on the plot.

To find the heart rate variability from this data, we took the difference between the heart rate data and the model in order to eliminate the recovery curve. We then found the standard deviation

² Chauhan, Vijay S., Andreq D. Krahn, Bruce D. Walker, George J. Klein, Allan C. Skanes, and Raymond Yee. "Dynamics During Exercise and Recovery in Healthy Subjects." *American Heart Journal* 144 (2002): 858-864.

of this difference. Subject 1's heart rate variability was 4.7 beats per minute, Subject 2's heart rate variability was 1.5 beats per minute, and Subject 3's heart rate variability was 2.6 beats per minute. Subject 1's R^2 value is lower than the other two because his heart rate variability was higher. All of these recovery heart rate variabilities are lower than the resting heart rate variabilities. An interesting further experiment would be to collect resting and recovery ECG data from many different subjects and compare the heart rate variabilities. This data could lead to interesting conclusions about vagal activity during exercise.

In this laboratory, we obtained raw ECG signals and found a good way to calculate heart rate. We did this by sampling our signal at a high enough rate, applying appropriate analog and digital filters to accentuate the QRS complexes, and then using a robust algorithm to detect QRS complexes. Using the information about heart rate that we obtained, we were able to study heart rate variability under different conditions.